CLAIMS:

1.   An encoding method for the compression of a video sequence divided in groups of frames decomposed by means of a three-dimensional (3D) wavelet transform leading to a given number of successive resolution levels corresponding to the decomposition levels of said transform, said method being based on a hierarchical subband encoding process

5   leading from the original set of picture elements (pixels) of each group of frames to transform coefficients constituting a hierarchical pyramid, and a spatio-temporal orientation tree - in which the roots are formed with the pixels of the approximation subband resulting from the 3D wavelet transform and the offspring of each of these pixels is formed with the pixels of the higher subbands corresponding to the image volume defined by these root pixels –

10   defining the spatio-temporal relationship inside said hierarchical pyramid, said encoding method being further characterized in that :

(A) the initial subband structure of the 3D wavelet transform is preserved by scanning the subbands one after the other in an order that respects the parent-offspring dependencies formed in said spatio-temporal tree ;

15   (B) flags " off / on " are added to each coefficient of the spatio-temporal tree in view of a progressive transmission of the most significant bits of the coefficients, these flags being such that at least one of them describes the state of a set of pixels and at least another one describes the state of a single pixel.

20   2.   An encoding method according to claim 1, characterized in that, for each bitplane, the tree scanning is spatially oriented, all the temporal resolutions being successively scanned inside each spatial scale and resolution flags being introduced between any two spatial scales.

25   3.   An encoding method according to claim 1, characterized in that, for each bitplane, the tree scanning is temporally oriented, all the spatial resolutions being successively scanned inside each temporal scale and resolution flags being introduced between any two temporal scales.

4.   An encoding method according to claim 1, characterized in that, for each bitplane, an intermediate tree scanning is performed, all the temporal and spatial resolutions of the same scale being jointly scanned and resolution flags being introduced between any two spatial/temporal scales.

5

5.   An encoding method according to anyone of claims 2 to 4, characterized in that two flags describe the state of a set of pixels and are, for each coefficient (x,y,z) of said spatio-temporal tree :

                - FS1 if $D(x,y,z)$ is still insignificant ;

10   - FS2 if $L(x,y,z)$ is still insignificant ;

where $D(x,y,z)$ is the set of coordinates of all the descendants of the node (x,y,z) and $L(x,y,z)$ = $D(x,y,z) - 0(x,y,z)$, with $0(x,y,z)$ being the set of coordinates of the direct offspring of the node (x,y,z), and two flags describe the state of a single pixel and

are :

15   - FP3 if the current pixel is significant ;

                - FP4 if it is not significant or if its significance is to be analyzed.

6.   An encoding method according to claim 5, characterized in that the exploration of the spatio-temporal tree, implemented in said scanning order, includes, after an

20   initialization step where the flag FP4 is put to all the coefficients of the lowest spatio-temporal subband and the flag FS1 to 7 over 8 coefficients of said lowest spatio-temporal subband, and the maximum significance level MSL is calculated, the following steps, carried out from the bitplane $n = MSL$ down to the bitplane $n = 0$ and from the lowest subband resolution down to the highest one :

25   (a) a first set of tests related to the set significance ;

                (1) if the flag FS1 is "on", then output $S_n (D(x,y,z))$:

                - if $S_n (D(x,y,z)) = 1$, then :

                                - for each (x',y',z') in $O(x,y,z)$, put flag FP4 ;

                                - remove flag FS1 from (x,y,z) ;

30   - if $L(x,y,z)$ not empty, then put flag FS2.

                (2) if flag FS2 is "on", then output $S_n (L(x,y,z))$:

                - if $S_n (L(x,y,z)) = 1$, then :

                                - for each (x',y',z') in $O(x,y,z)$, put flag FS1 ;

                                - remove flag FS2 from (x,y,z).

(b) a second set of tests related to the pixel significance :

    (1) if the flag FP3 is "on", then output = the n-th bit of (x,y,z) ;

    (2) if the flag FP4 is "on", then output $S_n(x,y,z)$:

        - if $S_n(x,y,z) = 1$, then :

                    - put flag FP3 "on" ;

                    - output sign (x,y,z) ;

                    - and remove flag FP4.

7.   An encoding method according to anyone of claims 1 to 6, characterized in that it also comprises a partial decoding step of the bitstream between two resolution flags, leading to a lower resolution/frame rate reconstructed video sequence.

8.   An encoding method according to claim 7, characterized in that the context used for the encoding of each bit related to the set significance in an arithmetic coding module is built using the bits of the same bitplane of the last scanned neighbouring wavelet coefficients in the same spatio-temporal subband, these bits being the bits output during the first set of tests related to the set significance.

9.   An encoding method according to claim 7, characterized in that the context used for the encoding of each bit related to the pixel significance in an arithmetic coding module is built using the bits of the same bitplane of the last scanned neighbouring wavelet coefficients in the same spatio-temporal subband, these bits being 1 if the neighbouring coefficients are marked by an FP3 flag and 0 if not.